**Biomedical Science 553  
Bioinformatics Applications II  
Course Syllabus**

**Instructor:** Randall Johnson, PhD

[johnson@hood.edu](mailto:johnson@hood.edu)

(240) 529-7363

Office: HT 159, Thurs 8 – 9 PM

**Course Description:** A continuation of Programming for Bioinformatics I. Statistics and R programming is extended to a broader range of biomedical applications. A deeper understanding will be gained of what can go wrong in data analysis, and best practices for reproducible research will be taught.

**Prerequisite:** BIFX 552 – Bioinformatics Applications I

**Required Textbook(s):** There is no required text for this course, but suggested readings will be given for students desiring to expand upon what is taught.

**Objectives:** BIFX 552 is meant to give students a solid foundation in statistical reasoning within the field of bioinformatics. In BIFX 553, students will expand their knowledge of statistical concepts to the analysis of primary data using intermediate to advanced programming skills. Tools for working as teams will also be explored, primarily through the use of the course git repository (see <https://github.com/johnsonra/BIFX553>).

**Student Learning Outcomes:** On completion of this course, students will be able to:

1. **Analyze** real data.
2. **Write readable R code** in the support of reproducible research.
3. **Present analysis results** in a clear, standard format.
4. **Work in small teams** to accomplish the objectives above.

**Homework:** Assignments will be given each week and will be due by the following Wednesday at 7 PM.

**Projects:** Two team projects will take place during the course. Teams should not include more than 4 people, and new teams will be established for each project.

**Grading:**

Homework 50%

Project I 25%  
Project II 25%

**Weather:** In the event of severe weather resulting in the closure of Hood College and the cancellation of a regularly scheduled class, an optional online discussion will be held during the regular class time. The online discussion will be recorded and made available to students for 1 week following the missed class, and a supplementary assignment may be given.

**Tentative Schedule of Course Topics**

|  |  |  |
| --- | --- | --- |
| **Week** | **Topic** | **Application** |
| Jan 19 | Pretest Evaluation R Tips and Tricks | Scripting and Programming |
| Jan 26 | Linear Regression Model building |  |
| Feb 2 | Generalized Linear Modeling I  Model fit and assumptions | Genetic Association |
| Feb 9 | Generalized Linear Modeling II  Complexity and Non-linearity | Genetic Association |
| Feb 16 | Survival Analysis  Kaplan Meier, Proportional Hazards | Genetic Association |
| Feb 23 | GWAS Considerations | GWAS |
| Mar 2 | Meta-Analysis | Micro Array |
| Mar 9 | Project I  Presentations & Review |  |
| Mar 16 | Spring Break |  |
| Mar 23 | Machine Learning I Supervised Learning | Prediction |
| Mar 30 | Machine Learning II Non-supervised Learning | Population Substructure |
| Apr 6 | Study Design I Bias, Variation and Error |  |
| Apr 13 | Study Design II Power and Sample Size Estimation |  |
| Apr 20 | Current Topics I | TBD |
| Apr 27 | Current Topics II | TBD |
| May 4 | Review |  |
| May 11 | Project II |  |